IMPROVED MUTUAL INFORMATION METHOD IN COMBINATION MODEL SELECTION FOR FORECASTING TOURIST ARRIVAL

MOHD ZULARIFFIN BIN MD MAAROF

A thesis submitted in fulfilment of the
requirements for the award of the degree
Doctor of Philosophy (Mathematics)

Faculty of Science
Universiti Teknologi Malaysia

FEBRUARY 2019

Dedicated to:

My beloved parents,

Md Maarof Mardi, Ramlah Abdul Latif

My supportive siblings,

Zulfadli, Noradilah, Nor Rislah, Zulkhairi, Zul Amin

My lovely wife,

Dr. Nur Syereena Nojumuddin

My dedicated Supervisor,

Prof. Datuk Dr. Zuhaimy Hj Ismail

My endless spirits

And all my friends.

This is for you.

ACKNOWLEDGEMENT

In the name of Allah, the Most Gracious and Most Merciful, all praise to Allah SWT, the Almighty, for His love has given me strength, perseverance, diligence and satisfaction in completing this project.

First and foremost, I would like to express my deepest gratitude, especially to my supervisor, Prof. Datuk Dr. Zuhaimy Hj. Ismail, who had taken a lot of effort to meticulously go through my work and come out with helpful suggestions. Not forgotten, million appreciations for my co-supervisor Dr. Mohamad Fadzli Ramli for their valuable critics and advices.

I would like to express my appreciation to the Ministry of Science, Technology and Innovation (MOSTI), SLAB/ SLAI program, and Research, University Grant (RUG) for supporting the scholarship along with my study.

Besides that, I also would like to acknowledge my special thanks to my supportive wife, Dr. Nur Syereena Nojumuddin for their suggestions, comment and moral support. Her effort is much appreciated. May Allah bless you.

Finally, I would like to express my greatest gratitude to my beloved family for their unstinting support and prayer. Without the family members support and prayer, this project would have been difficult at best. Thank You.

ABSTRACT

During the past several decades, a considerable amount of studies has been carried out on finding the highest accurate forecast model. Recently, it has been demonstrated that combining forecasts of individual models can improve forecast performance. Nevertheless, in practice, selecting individual forecast for model combination based on forecast accuracy evaluation might not have extracted all the significant information for the actual output forecast values. Hence, it is advocated to select the optimal individual model from theoretical and experimental aspects that may be able to offer more information to provide a better prediction of combination forecast model. Thus, the mutual information algorithm scaling proposed (MI-S-P) approach is proposed in this study to select the optimal individual model as an input for combination forecast model. Seven individual models and three linear combination methods are applied in this study to evaluate the effectiveness of the MI-S-P approach. The data used in this study is a short term 12 months ahead forecast which includes the monthly data on the top five international tourists arrival entering into Malaysia from the year 2000 to 2013. The results from this study is divided into two main parts, namely in-sample data (fitted model) and out-sample data (forecast model). The analyses show that the in-sample and out-sample values using MI-S-P model has successfully improve forecast accuracy on average by 2% compared to using all of individual forecast combination models. This study concludes that MI-S-P approach can be an alternative way in identifying the right optimal individual model for modelling combination forecast model.

ABSTRAK

Dalam beberapa dekad yang lalu, sejumlah besar kajian telah dijalankan untuk mencari model peramalan yang paling tepat. Baru-baru ini, menggabungkan model peramalan daripada beberapa model individu menunjukkan bahawa ianya boleh meningkatkan prestasi model peramalan. Walaubagaimanapun, secara praktikalnya, pemilihan model peramalan individu untuk membuat model gabungan berdasarkan penilaian ketepatan ramalan sahaja tidak mencukupi untuk mendapatkan maklumat penting bagi nilai data peramalan sebenar. Oleh itu, ianya adalah amat penting untuk memilih sub model individu yang optimum dari aspek teori dan eksperimen yang dapat memastikan lebih banyak informasi untuk menghasilkan model gabungan peramalan yang lebih baik. Oleh itu, pendekatan algoritma teori skala maklumat (MI-S-P) adalah dicadangkan dalam kajian ini untuk memilih sub model individu yang optimum sebagai input untuk menghasilkan model peramalan gabungan. Tujuh model individu dan tiga kaedah kombinasi model yang digunakan dalam kajian ini untuk menilai kecekapan model MI-S-P yang dicadangkan. Data yang digunakan dalam kajian ini adalah data bulanan jangka pendek 12 bulan ramalan ke hadapan iaitu daripada 5 negara pelancong terbanyak antarabangsa yang melawat Malaysia mulai tahun 2000 sehingga 2013. Hasil keputusan kajian ini dibahagikan kepada dua bahagian iaitu sampel data dalaman (sampel ujian) dan sampel data luaran (sampel ramalan). Analisis menunjukkan, pendekatan model MI-S-P bagi sampel ujian dan sampel ramalan berjaya memperbaiki ketepatan ramalan sebanyak 2% secara purata adalah lebih tepat berbanding menggabungkan semua model individu peramalan. Kesimpulannya, kajian ini menunjukkan bahawa pendekatan MI-S-P boleh menjadi pendekatan alternatif bagi mengenalpasti model individu optimum yang terbaik untuk menghasilkan model peramalan gabungan yang lebih tepat.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

e_t	-	The residual
p	-	The order of autoregressive model
q	-	The order of moving average model
P	-	The order of seasonal autoregressive model
Q	-	The order of seasonal moving average model
t	-	time
f_x	-	Value of fitness function
x	-	Number of chromosome
n	-	Number of observation in the time series
z_t	-	The number of monthly international tourist arrival
Z_t	-	The estimate numbers of international tourist arrival
H_0	-	Hypothesis one
H_1	-	Hypothesis two
S_1^2	-	Larger variance
S_2^2	-	Smaller variance
r	-	Correlation coefficient
$\overline{y}$	-	The mean of the time series
h	-	Maximum number of lag
k	-	the time lag
I	-	The difference of seasonal nor non-seasonal
ϕ_{kk}	-	Partial autocorrelation coefficient
ρ	-	population size
θ	-	Parameter for autoregressive model

ϕ	-	Parameter for moving average model
Φ	-	Parameter for seasonal moving average model
Θ	-	Parameter for seasonal moving average model
B	-	Backward shift operator
$\bar{x}_t$	-	The mean of difference time series data
x_t	-	The difference of time series data
$\frac{\partial y}{\partial \phi}$	-	Partial differential with respect to ϕ
$\frac{\partial y}{\partial \theta}$	-	Partial differential with respect to θ
w	-	weight of SA, VACO, DMSFE
ω	-	weight of SVRNN

LIST OF ABBREVIATIONS

ARIMA	-	Autoregressive integrated moving average
SARIMA	-	Seasonal autoregressive integrated moving average
AR	-	Autoregressive model
MA	-	Moving average model
ARMA	-	Autoregressive moving average model
GA	-	Genetic algorithm
BJ	-	Box Jenkins
GA-BJ	-	Genetic algorithm- Box Jenkins model
MSE	-	Mean square error
MAPE	-	Mean absolute percentage error
MAE	-	Mean absolute error
GA-SARIMA	-	Genetic algorithm-seasonal integrated moving average model
GA-ARIMA	-	Genetic algorithm- autoregressive integrated moving average model
SE	-	Standard error
ARFIMA	-	Autoregressive fractionally integrated moving average
ACF	-	Autocorrelation function
PACF	-	Partial autocorrelation function
FPE	-	Final prediction error
MEV	-	Minimum eigenvalue vector
MDL	-	Maximum distributed length
AIC	-	Akaike information criterion
AI	-	Artificial intelligence
ANN	-	Artificial neural network
VAR	-	Vector autoregressive model
ARDL	-	Autoregressive distributed lag

STSM	-	Structural time series model
TVP	-	Time varying parameter
GFS	-	Genetic fuzzy system
SA	-	Simple average
VACO	-	Variance covariance
DMSFE	-	Discounted mean square forecast error
WMES	-	Winter multiplicative exponential smoothing model
ES	-	Exponential smoothing
ESNA	-	Exponential smoothing no trend additive seasonality
ESNN	-	Exponential smoothing no trend no additive seasonality
ESNM	-	Exponential smoothing no trend multiplicative seasonality
SVRNN	-	Support vector regression neural network
SVR4	-	SVRNN with dimension four (D=4)
SVR5	-	SVRNN with dimension five (D=5)
SVR6	-	SVRNN with dimension six (D=6)
SVR7	-	SVRNN with dimension seven (D=7)
SVR8	-	SVRNN with dimension eight (D=8)
MI-S-P	-	Mutual information scaling proposed model
MI-S-SC	-	Mutual information scaling Shuang Cang model

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

This chapter provides an introduction of this study. The flow of this chapter starts with a background of the problem, statement of problem, research question, objective and scope of the study, research contribution, research data and research hypothesis. The organization of thesis plan in this study is includes at the end of the chapter as well.

1.2 Background of Problem

In forecasting fields, researchers continually explore for the most accurate individual model to generate a forecast. Much exertion has been committed in the course of recent decades to the advancement and evolution of forecasting models. Tourism forecasting model for time series data can be separated into two classes' namely linear method and nonlinear method. The most widely recognized of the time series linear forecasting model are naive I and naive II methods, the exponential smoothing (ES) and winter multiplicative exponential smoothing (WMES) model, the regression model, autoregressive integrated moving average model (ARIMA) and seasonal ARIMA (SARIMA). Among them, ARIMA and SARIMA models are the most developed gauging model that has been effectively tried in numerous practical applications.

Although linear model have been widely used in tourism studies, if the linear models neglected to perform well in both in-sample fitted model and out-sample forecasting model, more intricate nonlinear models ought to be considered. In light of this perspective, numerous researchers have additionally plough to nonlinear methods such as neural network (NN) model, genetic algorithm (GA) and support vector regression models (SVR) as an alternative to the development of tourism forecasting model. Indeed, even there are as yet a couple of questions about NN, GA, and SVR based tourism demand forecasting performance, it is for the most segment trusted that the nonlinear models (NN, SVR, GA and etc.) in modelling economic behaviour and efficiently helping wise decision making.

Some stationary phenomena practically speaking can be enamoured or if nothing else be approximated by linear and nonlinear models. Be that as it may, numerous nonstationary phenomena cannot be enamoured satisfactorily by these two linear and nonlinear models. For tourism contextual research, combination model is more precise to capture the tourism data for modelling forecast model. For this reason, an important motive to combine forecasts from different models is the fundamental assumption that one cannot identify the true process exactly, but different models may play a complementary role in the approximation of the data generating process.

Thus, more researchers focus on combination method either linear combination or nonlinear combination. The idea of consolidating combining model began with the original work 45 years back of Bates and Granger(Bates.J.M and Granger.C.W.J, 1969). Given two individual forecasts of a time series, they exhibit that a reasonable linear combination of the two forecasts may bring about a superior forecast results than the two original ones, in the feeling of a little error variance. . Previous studies have reported combination technique are likely to generate the best results which is higher than that of the individual models and leads to an improved forecast accuracy.

Hence in recent years, several studies have revealed that combination forecasting has been ended up being a profoundly effective in determining accurate forecasting model in numerous fields, which has been exhibited by observational studies(Song, Gao and Lin, 2013), (Andrawis, Atiya and El-Shishiny, 2011), (Kuan-

Yu and Chen, 2011), ,(Andrawis, Atiya and El-Shishiny, 2011; Li, Shi and Zhou, 2011), (Freitas and Rodrigues, 2006), (Shen, Li and Song, 2011).

However, the study of individual model process selection for combination forecast model is rarely studied in the literature. Much of the time, all available linear and nonlinear models are used as inputs for the combination model totally based on the result of the most accurate single model. This may lead to a wrong model selection in modelling combination forecast model. There is no comprehensive selection process in determining the optimum individual linear and nonlinear model in the sense of experimental and statistical theory perspective.

In view of this, while many researchers had studied combination forecasting between linear and nonlinear individual model, new process selection using information theory approach for modelling combination forecast model has received less attention among the researchers. Thus, this research will study a new algorithm model selection process for modelling combination forecast model using mutual information theory algorithm in the development of tourism forecasting model in Malaysia.

1.3 Statement of Problem

Numerous researchers demonstrate that combination model is showing greatly improved than individual model in terms of robustness and accuracy. Regularly, the best individual linear or nonlinear model is chosen from a few individual models that has higher accuracy. At that point, this best individual model will be combined with other best model in order to produce higher precision forecast model based on forecast error evaluation.

A noteworthy issue with this sort of combination model procedure is the point at which the total number of individual linear or nonlinear forecasting model is large. On the off chance that attempting every conceivable mix would include concentrated calculation and is to a great degree tedious. Shannon's information theory(Mackay

David, 2003) also argues that this procedure might not have extracted all the significant information for the actual output forecast values.

Thus, this study will investigate and develop a new procedure selection using improve mutual information theory algorithm to select the optimal individual linear or nonlinear model for a combination forecast model that could contains enough information to forecast the actual outputs.

1.4 Research Question

Questions arise when developing mutual information theory algorithm in modelling combination forecast model using optimal individual model from output of individual forecast model. It can be summarized as follows:

1. How to identify data pattern of time series data for modelling combination forecast model
2. How to design and develop forecast output value matrix using time series model including winter's multiplicative exponential smoothing method, support vector regression-neural network (SVR-NN), ARIMA and Seasonal ARIMA
3. How to select the optimal individual model using information theory for constructing a linear combination forecast model?
4. Is the accuracy and robustness of the information theory algorithm for modelling linear combination model produce a better prediction model compared to individual model?

1.5 Objective of the Study

The main objective of this study can be categorized into 4 parts. The objectives are stated as follows:

1. To determine data patterns of tourist arrival time series data using unit roots and stationary test
2. To generate forecast output values matrix of individual model for tourist arrival time series data
3. To identify the best optimal individual model of all individual models using information theory for linear combination model
4. To measure the robustness and effectiveness of the proposed model of information theory for short term forecast data.

1.6 Limitation of the Study

The limits of this study are:

i. This study focuses on modeling linear combination model (simple average, Variance-Covariance, and Discounted mean square forecast error), individual time series model (ARIMA,SARIMA and Exponential smoothing), non-linear combination model (support vector regression - neural network).

ii. Forecast accuracy in this study will be defined by measuring the lowest error in term of mean absolute scale error (MASE) and mean absolute percentage error (MAPE).

iii. Mutual information theory is applied to choose the optimal individual model as an input for linear combination model.

iv. The data used is the secondary data of tourist arrival to Malaysia from the period of year 1999 to 2013 as a case study to assess the adequacy of the proposed forecast model.

v. The econometric forecast model using gross domestic product (GDP) and consumer price index (CPI) variable is not covered in this study because this study focus on developing tourism forecast model not on the impact of tourism analysis .

vi. The forecast horizon of this study is limited to short term forecast for 12 months data ahead only.

1.7 Research Contribution

Although many researchers in previous literatures such as research conducted by Bates and Granger (1969), Zhang (2003), Lessmann et al (2012) and Wang and Hu (2015) had been studied in many mathematical combinations forecast model development in finding the accurate one, there are as yet still no accurate combination forecast model which are able to determine time series data for modelling tourist arrival forecast model. This study attempts to determine the significant procedure in finding the best combination forecast model for forecasting tourist arrival time series data to Malaysia. The expected contribution of this study is five.

First, the guidelines and procedures for identifying the data pattern analysis before model development is identified are presented. This guidelines will be useful for the purpose of this current research as well as for those who conducting similar study. This guideline also presents the method and ways to determine the data pattern characteristics of time series data. The details of it can be found in Chapter 3 and 4.

Second, this study presents how to develop a linear and nonlinear fitted and forecast model in modelling accurate individual model using seasonal and nonseasonal time series data. Chapter 5 shows the details on developing fitted and forecast model for linear and nonlinear model. The fitted and forecast output of individual model is a foundation study for the development of combination forecast model.

Third, this study attempts to develop a new procedure for modelling combination forecast model using mutual information theory algorithm as the selection tools. This study presents an evidence to prove an individual forecasting model is not always best in all cases in developing accurate forecast model. Chapter 6 shows the theoretical and experimental work of the development mutual information theory algorithm gives higher accuracy than the individual forecast model itself.

Fourth, the modified selection procedure of using mutual information theory algorithm gives higher forecast accuracy as this study compares with the existing methodology using the same data input. This is due to the new algorithm may be able to offer more information to provide a better prediction model. The details of theoretical framework and results analysis of the selection procedure of using mutual information theory algorithm can be found in Chapter 3 and 6.

Last but not least, the contribution of this study is the development of a system for selecting individual model procedure for modelling combination forecast model. This procedure has been develop the coding and VBA interface and it is not available in any of the current statistical packages. The development of this coding is vital not only for the tourism forecast model only, but it can also be applied for any forecasting model development in any sector. The programs of this system user can automatically choose the parameter and calculate the output. User do not have to know the algorithm behind the program as well. The development of the coding can be found in Appendix C.

1.8 Research Data

Five different types of data are used in this work. The time series data are top 5 countries (Singapore, Indonesia, Brunei, Thailand, and China) monthly tourist arrival to Malaysia from year 2000 until 2013. The data is secondary data provided by Malaysian Tourism Promotion Board.

REFERENCES

Abdel-salam, E. A. and Al-muhiameed, Z. I. A. (2015) 'Analytic Solutions of the Space-Time Fractional Combined KdV-mKdV Equation', 2015.

Amir, S. *et al.* (2015) 'Understanding domestic and international t ourists ' expenditure pattern in Melaka , Malaysia : result of CHAID analysis', *Procedia - Social and Behavioral Sciences*. Elsevier B.V., 172, pp. 390–397. doi: 10.1016/j.sbspro.2015.01.386.

Andrawis, R. R., Atiya, A. F. and El-Shishiny, H. (2011) 'Combination of long term and short term forecasts, with application to tourism demand forecasting', *International Journal of Forecasting*. Elsevier B.V., 27(3), pp. 870–886. doi: 10.1016/j.ijforecast.2010.05.019.

Athanasopoulos, G. *et al.* (2011) 'The tourism forecasting competition', *International Journal of Forecasting*. Elsevier B.V., 27(3), pp. 822–844. doi: 10.1016/j.ijforecast.2010.04.009.

Athanasopoulos, G. and Hyndman, R. J. (2008) 'Modelling and forecasting Australian domestic tourism', *Tourism Management*, 29(1), pp. 19–31. doi: 10.1016/j.tourman.2007.04.009.

Baba.C and Kisinbay.T (2011) *Predicting recessions: A new approach for identifying leading indicators and forecast combinations.*

Baba, S. and Matsuishi, T. (2014) 'Evaluation of the predictability of fishing forecasts using information theory', *Fisheries Science*, 80(3), pp. 427–434. doi: 10.1007/s12562-014-0736-8.

Bates.J.M and Granger.C.W.J (1969) 'The Combination of Forecasts', *Operational Research Society*, 20(4), pp. 451–468.

Battiti, R. (1994) 'Using mutual information for selecting features in supervised neural net learning.', *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 5(4), pp. 537–50. doi: 10.1109/72.298224.

Bermúdez, J. D., Segura, J. V. and Vercher, E. (2007) 'Holt–Winters

Forecasting: An Alternative Formulation Applied to UK Air Passenger Data', *Journal of Applied Statistics*, 34(9), pp. 1075–1090. doi: 10.1080/02664760701592125.

Boudi, A. K. (2011) 'Time Series Modeling of Tourism in Southwest Algeria Case Study Bechar as Tourist Destination', 69(69).

Box, G. E. . and Jenkins, G. (1970) *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco CA.

Brown, R. G. (1959) *Statistical forecasting for inventory control*. New York: Mcgrow-Hill.

Burger, C. J. S. . *et al.* (2001) 'A practitioners guide to time-series methods for tourism demand forecasting — a case study of Durban, South Africa', *Tourism Management*, 22(4), pp. 403–409. doi: 10.1016/S0261-5177(00)00068-6.

Van Calster, T., Baesens, B. and Lemahieu, W. (2017) 'ProfARIMA: A profit-driven order identification algorithm for ARIMA models in sales forecasting', *Applied Soft Computing Journal*. Elsevier B.V., 60, pp. 775–785. doi: 10.1016/j.asoc.2017.02.011.

Cang, S. (2013) 'A Comparative Analysis of Three Types of Tourism Demand Forecasting Models : Individual , Linear Combination and Non-linear Combination'. doi: 10.1002/jtr.

Cang, S. and Yu, H. (2012) 'Mutual information based input feature selection for classification problems', *Decision Support Systems*. Elsevier B.V., 54(1), pp. 691–698. doi: 10.1016/j.dss.2012.08.014.

Cang, S. and Yu, H. (2013) 'A combination selection algorithm on forecasting', *European Journal of Operational Research*. Elsevier B.V., 234(August 2013), pp. 127–139. doi: 10.1016/j.ejor.2013.08.045.

Cao, X. H., Stojkovic, I. and Obradovic, Z. (2016) 'A robust data scaling algorithm to improve classification accuracies in biomedical data', *BMC Bioinformatics*. BMC Bioinformatics, 17(1), p. 359. doi: 10.1186/s12859-016-1236-x.

Carbonneau, R., Laframboise, K. and Vahidov, R. (2008) 'Application of machine learning techniques for supply chain demand forecasting', *European Journal of Operational Research*, 184(3), pp. 1140–1154. doi: 10.1016/j.ejor.2006.12.004.

Chatfield, C. (2001) *Time Series Forecasting*. Chapman & Hall CRC FLorida.

Che, J. (2015) 'Optimal sub-models selection algorithm for combination forecasting model', *Neurocomputing*. Elsevier, 151, pp. 364–375. doi:

10.1016/j.neucom.2014.09.028.

Chen, C.-F., Lai, M.-C. and Yeh, C.-C. (2012) ‘Forecasting tourism demand based on empirical mode decomposition and neural network’, *Knowledge-Based Systems*. Elsevier B.V., 26(2012), pp. 281–287. doi: 10.1016/j.knosys.2011.09.002.

Chen, K.-Y. (2011a) ‘Combining linear and nonlinear model in forecasting tourism demand’, *Expert Systems with Applications*. Elsevier Ltd, 38(8), pp. 10368–10376. doi: 10.1016/j.eswa.2011.02.049.

Chen, K.-Y. (2011b) ‘Combining linear and nonlinear model in forecasting tourism demand’, *Expert Systems with Applications*. Elsevier Ltd, 38(8), pp. 10368–10376. doi: 10.1016/j.eswa.2011.02.049.

Cho, V. (2003) ‘A comparison of three different approaches to tourist arrival forecasting’, *Tourism Management*, 24(3), pp. 323–330. doi: 10.1016/S0261-5177(02)00068-7.

Chu, F.-L. (1998) ‘Forecasting tourism: a combined approach’, *Tourism Management*, 19(6), pp. 515–520. doi: 10.1016/S0261-5177(98)00053-3.

Claveria, O., Monte, E. and Torra, S. (2016) ‘Combination forecasts of tourism demand with machine learning models’, *Applied Economics Letters*, 23, pp. 428–431.

Clemen, T. . (1989) ‘Combining forecasts : A review and annotated bibliography’, *International Journal of Forecasting*, 5, pp. 559–583.

Clements, M. (2004) ‘Evaluating Journal of Forecasting’, *Economic Journal*, 114, pp. 844–866.

Costantini, M. and Pappalardo, C. (2010) ‘A hierarchical procedure for the combination of forecasts’, *International Journal of Forecasting*. Elsevier B.V., 26(4), pp. 725–743. doi: 10.1016/j.ijforecast.2009.09.006.

Cover, T. M. and Thomas, J. A. (2005) *Elements of Information Theory*. 2nd editio. John Wiley & Sons, Inc.

Dickey, D. A. and Fuller, W. A. (1981) ‘LIKELIHOOD RATIO STATISTICS FOR AUTOREGRESSIVE TIME SERIES WITH A UNIT’, *Econometrica*, 49(4), pp. 1057–1072.

Diebold.F.X and Pauly.P (1987) ‘Structural change and the combination of forecasts’, *Journal of Forecasting*, 6, pp. 21–40.

Estévez, P. a *et al.* (2009) ‘Normalized mutual information feature selection.’, *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 20(2), pp. 189–201. doi: 10.1109/TNN.2008.2005601.

Freitas, P. S. a. and Rodrigues, A. J. L. (2006) 'Model combination in neural-based forecasting', *European Journal of Operational Research*, 173(3), pp. 801–814. doi: 10.1016/j.ejor.2005.06.057.

Hamilton and Douglas, J. (1994) *Time Series Analysis*. Princeton: Princeton university oress.

Hanke, J. E. and Wichern, D. W. (2005) *Business Forecasting*. Eight Edit. Pearson Prentice Hall.

Holt, C. C. (2004) 'Forecasting seasonals and trends by exponentially weighted moving averages', *International Journal of Forecasting*, 20(1), pp. 5–10. doi: 10.1016/j.ijforecast.2003.09.015.

Hsu, C.-I. *et al.* (2009) 'Predicting tourism loyalty using an integrated Bayesian network mechanism', *Expert Systems with Applications*. Elsevier Ltd, 36(9), pp. 11760–11763. doi: 10.1016/j.eswa.2009.04.010.

Ibrahim, Y. (2010) 'Forecasting International Tourism Demand in Malaysia Using Box Jenkins Sarima Application', 3(2).

Ismail, Z. and Zulariffin, M. (no date) 'Alteration of Box-Jenkins Methodology by Implementing'.

Journal, S., Statistical, R. and Series, S. (2014) 'Experience with Forecasting Univariate Time Series and the Combination of Forecasts Author (s): P . Newbold and C . W . J . Granger Experience with Forecasting Univariate Time Series and the Combination of Forecasts', 137(2), pp. 131–165.

Jun, W. *et al.* (2018) 'Modeling a combined forecast algorithm based on sequence patterns and near characteristics: An application for tourism demand forecasting', *Chaos, Solitons and Fractals*. Elsevier Ltd, 108, pp. 136–147. doi: 10.1016/j.chaos.2018.01.028.

Kisinbay, T. (2010) 'The Use of Encompassing Tests for Forecast Combinations', *Journal of Forecasting*, 29(2010), pp. 715–727.

Kuan-Yu and Chen (2011) 'Combining linear and nonlinear model in forecasting tourism demand', *Expert Systems with Applications*, 38, pp. 10368–10376. doi: 10.1016/j.eswa.2011.02.049.

Kwak, N. and Choi, C.-H. (2002) 'Input feature selection for classification problems.', *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 13(1), pp. 143–59. doi: 10.1109/72.977291.

Lessmann, S. *et al.* (2012) 'A new methodology for generating and combining statistical forecasting models to enhance competitive event prediction', *European Journal of Operational Research*. Elsevier B.V., 218(1), pp. 163–174. doi: 10.1016/j.ejor.2011.10.032.

Li, G. *et al.* (2006) 'Tourism Demand Forecasting: A Time Varying Parameter Error Correction Model', *Journal of Travel Research*, 45(2), pp. 175–185. doi: 10.1177/0047287506291596.

Li, G., Shi, J. and Zhou, J. (2011) 'Bayesian adaptive combination of short-term wind speed forecasts from neural network models', *Renewable Energy*. Elsevier Ltd, 36(1), pp. 352–359. doi: 10.1016/j.renene.2010.06.049.

Liang, H., Sun, X. and Sun, Y. (2017) 'Text feature extraction based on deep learning : a review'. EURASIP Journal on Wireless Communications and Networking, pp. 1–12. doi: 10.1186/s13638-017-0993-1.

Lim, C. and Mcaleer, M. (2001) 'Asian Tourism to Australia', 28(1), pp. 68–82.

Lin, C. J., Chen, H. F. and Lee, T. S. (2011) 'Forecasting Tourism Demand Using Time Series, Artificial Neural Networks and Multivariate Adaptive Regression Splines:Evidence from Taiwan', *International Journal of Business Administration*, 2(2). doi: 10.5430/ijba.v2n2p14.

M.Toshinori (1998) *Fundamentals of the new artificial intelligence*. New York: Springer.

Mackay David, J. . (2003) *Information theory, inference, and learning algorithms*. Cambridge University Press.

Mansor, N. *et al.* (2015) 'Agro Tourism Potential in Malaysia', 1(2), pp. 37–44.

McCleary, R., Hay, R. and McDowell, D. (1980) *Applied time series analysis for the social sciences*. Sage, Los Angeles.

Md Maarof, M. Z., Ismail, Z. and Fadzli, M. (2014) 'Optimization of SARIMA model using genetic algorithm method in forecasting Singapore tourist arrivals to Malaysia', *Applied Mathematical Sciences*, 8(169–172). doi: 10.12988/ams.2014.410847.

Naude, W. and Saayman, A. (2005) 'Determinants of tourist arrivals in Africa: a panel data regression analysis', *Tourism Economics*, 11, pp. 365–391.

Ni, K. S. and Nguyen, T. Q. (2007) 'Image superresolution using support vector regression.', *IEEE transactions on image processing : a publication of the IEEE*

Signal Processing Society, 16(6), pp. 1596–610. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23757596.

Ong, C.-S., Huang, J.-J. and Tzeng, G.-H. (2005) 'Model identification of ARIMA family using genetic algorithms', *Applied Mathematics and Computation*, 164(3), pp. 885–912. doi: 10.1016/j.amc.2004.06.044.

Page, S., Song, H. and Wu, D. C. (2011) 'Assessing the Impacts of the Global Economic Crisis and Swine Flu on Inbound Tourism Demand in the United Kingdom', *Journal of Travel Research*, 51(2), pp. 142–153. doi: 10.1177/0047287511400754.

Pai, P., Hung, K. . and Lin, K. . (2014) 'Tourism demand forecasting using novel hybrid system', *Expert Systems with Applications*, 41(8), pp. 3691–3702.

Papana, A. and Kugiumtzis, D. (2009) 'Evaluation of mutual information estimators for time series', *International journal of bifurcation and chaos*, 19(12), pp. 111–121.

Peng, B., Song, H. and Crouch, G. I. (2014) 'A meta-analysis of international tourism demand forecasting and implications for practice', *Tourism Management*. Pergamon, 45, pp. 181–193. doi: 10.1016/J.TOURMAN.2014.04.005.

Peng, H., Long, F. and Ding, C. (2005) 'Feature selection based on mutual information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy', *IEEE Transactions on pattern analysis and machine intelligence2*, 27(8), pp. 1226–1238. Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Feature+Selection+Based+on+Mutual+Information+:#4 (Accessed: 29 November 2014).

Petropoulos, F. and Kourentzes, N. (2014) 'Forecast combinations for intermittent demand'. Nature Publishing Group, 66(6), pp. 914–924. doi: 10.1057/jors.2014.62.

Phillips, P. C. . and Perron, P. (1988) 'Testing for a unit root in time series regression', *Biometrika*, 75(2), pp. 335–346.

Rossi, F., Lendasse, A. and Franc, D. (2006) 'Mutual information for the selection of relevant variables in spectrometric nonlinear modelling B', 80, pp. 215–226. doi: 10.1016/j.chemolab.2005.06.010.

Shahrabi, J., Hadavandi, E. and Asadi, S. (2013) 'Developing a hybrid intelligent model for forecasting problems: Case study of tourism demand time series', *Knowledge-Based Systems*, 43(2013), pp. 112–122. doi: 10.1016/j.knosys.2013.01.014.

Shen, S., Li, G. and Song, H. (2008a) 'An Assessment of Combining Tourism Demand', (July). doi: 10.1177/0047287508321199.

Shen, S., Li, G. and Song, H. (2008b) 'An Assessment of Combining Tourism Demand Forecasts over Different Time Horizons', *Journal of Travel Research*, 47(2), pp. 197–207. doi: 10.1177/0047287508321199.

Shen, S., Li, G. and Song, H. (2011) 'Combination forecasts of International tourism demand', *Annals of Tourism Research*. Elsevier Ltd, 38(1), pp. 72–89. doi: 10.1016/j.annals.2010.05.003.

Shitan, M. (2008) 'Time Series Modelling of Tourist Arrivals to Malaysia', *interstat.statjournals.net*, pp. 1–12. Available at: http://interstat.statjournals.net/YEAR/2008/articles/0810005.pdf (Accessed: 9 June 2011).

Song, H., Gao, B. Z. and Lin, V. S. (2013) 'Combining statistical and judgmental forecasts via a web-based tourism demand forecasting system', *International Journal of Forecasting*. Elsevier B.V., 29(2), pp. 295–310. doi: 10.1016/j.ijforecast.2011.12.003.

Song, H. and Li, G. (2008) 'Tourism demand modelling and forecasting—A review of recent research', *Tourism Management*, 29(2), pp. 203–220. doi: 10.1016/j.tourman.2007.07.016.

Sridhar, D. V., Bartlett, E. B. and Seagrave, R. C. (1999) 'An information theoretic approach for combining neural network process models', *Neural Networks*, 12(6), pp. 915–926. doi: 10.1016/S0893-6080(99)00030-1.

Suhartono (2011) 'Time Series Forecasting by using Seasonal Autoregressive Integrated Moving Average: Subset, Multiplicative or Additive Model', *Journal of Mathematics and Statistics*, pp. 20–27. doi: 10.3844/jmssp.2011.20.27.

Tay, F. E. H. and Cao, L. (2001) 'Application of support vector machines in ÿnancial time series forecasting', 29, pp. 309–317.

Tran, H. D., Muttil, N. and Perera, B. J. C. (2015) 'Selection of significant input variables for time series forecasting', *Environmental Modelling and Software*. Elsevier Ltd, 64, pp. 156–163. doi: 10.1016/j.envsoft.2014.11.018.

Vapnik, V. (1995) *The nature of statistical learning theory*. Springer New York.

Vapnik, V., S.Golowich and A.Smola (1996) 'Support vector machine for function approximation regression estimation, and signal processing', *Advances in*

Neural Information Processing Systems, 9, pp. 281–287.

Wang, H. and Zhao, W. (2009) 'ARIMA Model Estimated by Particle Swarm Optimization Algorithm for Consumer Price Index Forecasting', pp. 48–58.

Wang, J. and Hu, J. (2015) 'A robust combination approach for short-term wind speed forecasting and analysis e Combination of the ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and LSSVM (Least Square SVM) forec', *Energy*. Elsevier Ltd, 93, pp. 41–56. doi: 10.1016/j.energy.2015.08.045.

Winkler.R.L and Makridakis.S (1983) 'The Combination of Forecasts', *Journal of the Royal Statistical Society, Series A*, 146(2), pp. 150–157.

Wong, K. K. F. *et al.* (2007) 'Tourism forecasting: To combine or not to combine?', *Tourism Management*, 28(4), pp. 1068–1078. doi: 10.1016/j.tourman.2006.08.003.

Xu, X. *et al.* (2016) 'Forecasting tourism demand by extracting fuzzy Takagi–Sugeno rules from trained SVMs', *CAAI Transactions on Intelligence Technology*. Elsevier Ltd, 1(1), pp. 30–42. doi: 10.1016/j.trit.2016.03.004.

Zhang, G. P. (2003) 'Time series forecasting using a hybrid ARIMA and neural network model', *Neurocomputing*, 50, pp. 159–175.